

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-108844  
 (43)Date of publication of application : 12.04.2002

(51)Int.CI. G06F 17/21  
 G06F 12/00

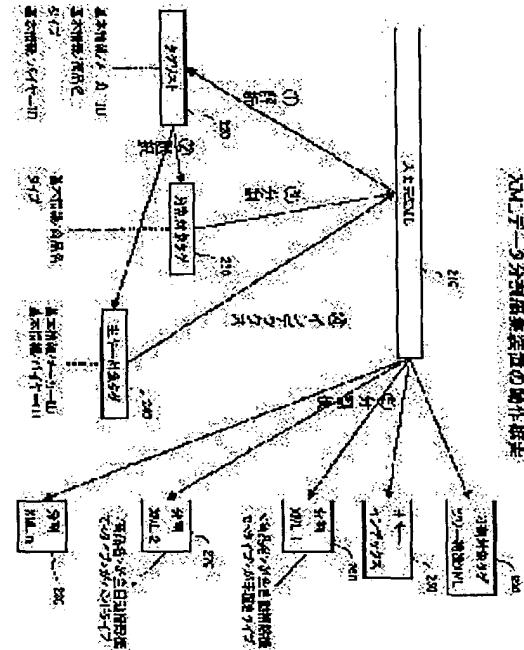
(21)Application number : 2000-295352 (71)Applicant : HITACHI SOFTWARE ENG CO LTD  
 (22)Date of filing : 28.09.2000 (72)Inventor : KIJIMA NORIKAZU

## (54) XML DATA DIVISION EDITING DEVICE

## (57)Abstract:

PROBLEM TO BE SOLVED: To quickly retrieve and efficiently edit a large quantity of XML data by dividing them.

SOLUTION: Input source XML data 110 are analyzed to prepare a tag list 120. A main key index 150 for corresponding a main key to the split XML data, split XMLs 160, 170 and 180 divided by a split object tag and a split object tag tree structure XML 190 are prepared from the tag list 120 by using the split object tag 130, a main key object tag 140 and the data 110.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

## (12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2002-108844  
(P2002-108844A)  
(43)公開日 平成14年4月12日(2002.4.12)

(51)Int. C.7  
G 06 F 17/21  
12/00  
5 0 1  
5 7 0  
5 4 7

発明記号  
P 1  
G 06 F  
17/21  
12/00  
5 0 1  
5 7 0  
5 4 7

発明記号  
P 1  
G 06 F  
17/21  
12/00  
5 0 1  
5 7 0  
5 4 7

発明記号  
P 1  
G 06 F  
17/21  
12/00  
5 0 1  
5 7 0  
5 4 7

審査請求未請求 請求項の数4

OJ (全13頁)

(21)出願番号 特願2000-295352 (P2000-295352)  
(22)出願日 平成12年9月28日(2000.9.28)

(71)出願人 日立ソフトウェアエンジニアリング株式会社  
神奈川県横浜市中区尾上町6丁目81番地  
(72)発明者 木島 敏和  
神奈川県横浜市中区尾上町6丁目81番地  
日立ソフトウェアエンジニアリング株式会社  
社内  
(74)代理人 100073760  
弁理士 鈴木 誠  
Fターム(参考) 5B009 Q111 S403 V402  
5B082 E401 G408

(71)出願人 0002330565

(72)発明者 木島 敏和

(74)代理人 100073760

Fターム(参考) 5B009 Q111 S403 V402  
5B082 E401 G408

【特許請求の範囲】 XMLデータを編集・検索する装置であ

る、肥大化したXMLデータを効率的に扱うためにはX

MLデータを分割する必要がある。

【請求項1】 XMLデータのタグとタグの間に基づいて、入力元XMLデータを解析してタグリストを生成する手段と、前記タグリストから選択された主キーが象

する手段と、前記タグリストを用いて前記XMLデータを分割するタグと分割対象タグを用いて前記XMLデータを分割する手段とを有することを特徴とするXMLデータ分割装置。

【請求項2】 請求項1記載のXMLデータ分割装置において、入力元XMLデータについて、1レコード10の値の出現頻度などに基づいて解析してタグリストを生

成する手段と、該タグリストから選択した主キーが象タグと分割対象タグを用いて入力元XMLデータを分割する手段を備えることを主な特徴とする。

【請求項3】 本発明では、XMLデータ中に含まれるタ

値を持ち、1レコード中の開閉層に1つだけ含まれるタグを対象にして、同じタグ値を取る同名タグとタグの出

現頻度を計測し、タグリストを生成する。このタグリストから選択したタグに基づいて、ソリューションのディレクトリを作成し、複数のXMLファイルへ分割する。

【請求項4】 請求項3記載のXMLデータ分割装置において、主キーと分割XMLファイルを対応づける構造を作成するXMLファイルを作成する。

【発明の実施の形態】 以下、図面により本発明の一実施形態の概要について説明する。図1は、本発明によるXMLデータ分割装置の構成図である。

本XMLデータ分割装置は、処理装置(CPU)10、表示装置20、キーボード30、マウス40、モディ

ム50、RAMなどの記憶装置60、及びハードディスクなどの外記憶装置70などから構成されるが、このハ

ードウェア構成自体は所要パソコンやワークステーションなどと基本的に同様である。ここで、処理装置10は

本発明に係する手段としてタグリスト生成手段11とXMLデータ分割装置12を同様である。ここで、処理装置10はタグリスト120を生成する手段としてタグリスト生成手段11とXMLデータ分割装置12を同様である。処理装置10はタグリスト120、分割対象タグ130、主キー対象タグ140を一時的に格納する。外部記憶装置70

は、あらかじめ取得した入力元XMLデータ110に加え、処理装置10の処理結果としての主キーインデックス150、複数の分割XMLデータ160、170、180、及び分割対象タグソリューションXML190などを格納する。

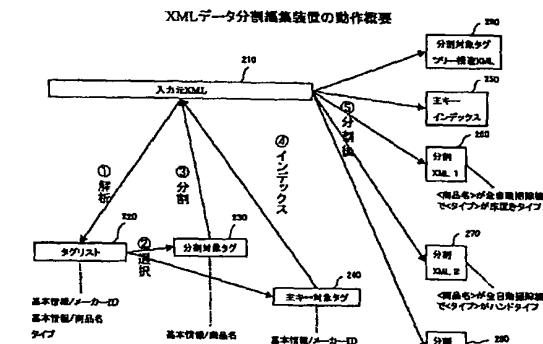
【0009】 図2は本XMLデータ分割装置の動作概要を示した図である。処理装置10のタグリスト生成手段11は、外部記憶装置70から入力元XMLデータ

110を取り込み、該入力元XMLデータ110をタグの出現頻度で解析してタグリスト120を生成し、記憶装置120に格納するとともに表示装置20に表示する。

ユーパーは、キーボード30やマウス40により、表示装置20に表示されたタグリスト120から分割対象

【0010】 【発明が解決しようとする課題】 従来技術では、XMLデータとして使用する場合、データ量が少ない場合は問題ないが、データ量が肥大化した場合、検索、更新等に非常に時間がかかる上、且で確認することが困難である。

【0011】 【発明の詳細な説明】



## (4)【発明の名称】 XMLデータ分割装置

【0012】 【要約】 大量なXMLデータを分割することによって、迅速に検索し、効率的に編集できるようにする。  
【解決手段】 入力元XMLデータ110を解してタグリスト120を作成し、タグリスト120から、分割対象タグ130と主キー対象タグ140と入力元XMLデータ110を抽出して、主キーインデックス150、分割XMLデータ160、170、180、分割対象タグソリューションXML190を作成する。

【0013】 従来、XMLデータをすべて読み込み、検索、編集する方法においては、スタイルシートやクリエーターの読み込み込みが非常に困難となる。一方で、一つのXMLデータ内に複数のXMLデータのみ対象としており、一つのXMLデータであってもすべてのXMLデータを読み込む必要があり、XMLデータを効率的に編集するることは困難である。

【0014】 【発明が解決しようとする課題】 従来技術では、XMLデータとして使用する場合、データ量が少ない場合は問題ないが、データ量が肥大化した場合、検索、更新等に非常に時間がかかる上、且で確認することが困難である。

【0015】 本発明の目的は、上記従来技術の問題点を解消し、大規模なXMLデータの効率的な編集、検索等を実現するXMLデータ分割装置を提供することに

ある。肥大化したXMLデータを効率的に扱うためにはX

MLデータを分割する必要がある。

【0016】 本発明の目的は、上記目的を達成するためには、入力元XMLデータを、そのタグとタグの出現頻度などに基づいて解析してタグリストを生

成する手段と、該タグリストから選択した主キーが象タグと分割対象タグを用いて入力元XMLデータを分割する手段を備えることを主な特徴とする。

【0017】 本発明では、XMLデータ中に含まれるタ

値を持ち、1レコード中の開閉層に1つだけ含まれるタ

グを対象にして、同じタグ値を取る同名タグとタグの出

現頻度を計測し、タグリストを生成する。このタグリストから選択したタグに基づいて、ソリューションのディレ

クトリを作成し、複数のXMLファイルへ分割する。

【0018】 本発明では、XMLデータ分割装置においては更新をすれば

レコード上の主キーとなる項目については更新をすれば

XMLデータを対応づけるXMLファイルを作成する。

【0019】 【発明の実施の形態】 以下、図面により本発明の一実施形態の構成図である。

【0020】 【発明の詳細な説明】

【0021】 【発明の詳細な説明】 本発明はXMLデータ分割装置

に係り、特に大規模なXMLデータの編集、検索などに好適なXMLデータ分割装置に属する。

【0022】 【発明の詳細な説明】 本発明は所要パソコンやワークステーションなどと基本構成は同様であるが、このハ

ードウェア構成自体は所要RAM60、及びハードディスク70などから構成されるが、このハ

ードウェア構成自体は所要パソコンやワークステーションなどと基本構成は同様である。ここで、処理装置10は

本発明に係する手段としてタグリスト生成手段11と

XMLデータ分割装置12をすべて読み込まなければ編集す

ることができない。また、大量のXMLデータの中から

目的のデータを見つけることが困難である。このような

理由から、XMLデータ量が少ない場合はXMLデータを

編集、検索することが難しくないが、XMLデータ量

が増えると、XMLデータを編集することが非常に困難

となる。

【0023】 従来、XMLデータをすべて読み込み、検

索、編集する方法においては、スタイルシートやクリエ

ーターの読み込み込みが非常に困難となる。一方で、

一つのXMLデータ内に複数のXMLデータであってもすべてのXMLデータを読み込む必要があり、XML

データを効率的に編集するためには、XMLデータを

効率的に分割する手段が必要である。そこで、本発明

では、主キーとタグリストを用いてXMLデータを分割す

ることとする。主キーとタグリストを用いてXMLデータを分割する手段としては別個成り立つ。

【0024】 図2は本XMLデータ分割装置の動作概要を示した図である。処理装置10のタグリスト生成

手段11は、外部記憶装置70から入力元XMLデータ

110を取り込み、該入力元XMLデータ110をタグ

の出現頻度で解析してタグリスト120を生成し、記憶

装置120に格納するとともに表示装置20に表示す

る。ユーザーは、キーボード30やマウス40により、表

示装置20に表示されたタグリスト120から分割対象

【0025】 【発明の詳細な説明】

【0026】 【発明の詳細な説明】

【0027】 【発明の詳細な説明】

【0028】 【発明の詳細な説明】

【0029】 【発明の詳細な説明】

【0030】 【発明の詳細な説明】

【0031】 【発明の詳細な説明】

【0032】 【発明の詳細な説明】

【0033】 【発明の詳細な説明】

3

6

タグ7 1 3 0と主キー対象タグ1 4 0を選択する。この分割対象タグリーフィルタ XML 1 3 0と主キー対象タグ1 4 0は記憶装置6 0に格納される。処理装置1 0のXMLデータ分割手段1 2は、入力元XMLデータ1 1 0を、分割対象タグ1 3 0によって分割して分割XMLデータ1 6 0、1 7 0、1 8 0を生成し、同時に主キー対象タグ1 4 0の値と該当レコードが取り出される。そこで、一つの商品データ(1レコード)7 2 0に着目し(ステップ3 0 3)、その中にまだ取り出していないタグがまだ存在するか判定する(ステップ3 0 4)。1商品データ7 0の値にはまだ取り出していないタグが存在する。そこで、一つの商品データ(1レコード)7 2 1を取り出される。ステップ3 0 5)。そして、この取り出したタグは値を持ち、同階層タグに同名タグが存在するか判定する(ステップ3 0 6)。「基本情報」タグ7 2 1は値を持たない。そこで、「基本情報」タグ7 2 1を取り出される。ステップ3 0 7)。その間にまだ取り出していないタグが存在するか判定する(ステップ3 0 8)。

【0 0 1 8】図4はタグリストの具体例であり、図7では、一つの商品データ(1レコード)7 2 0に着目する。そこで、一つの商品データ(1レコード)7 2 0に着目し(ステップ5 0 5)。そして、複数するタグが存在する(ステップ5 0 6)。「基本情報」タグ7 2 1は主キー対象タグではない。そこで、次に着目するタグを同じ値の多い順とタグの値を並べたものである。図4中、4 1 0がタグリストを示す。該データリスト4 1 0は、入力XMLデータ7 2 1中に含まれる11レコード(1商品データ)中に一つだけ出現するタグ名4 1 2が入力XMLデータ7 1 0中に含まれる同名タグ数4 1 3、同タグ名で同じタグの値の数4 1 4の大さい順に順位4 1 1付けして並べリストである。ここでは、同タグ名で同じタグの値の数4 1 4は同名タグが同じ値をとる数を大きい順に数器目まで示し、タグリスト4 1 0では4番目まで出力されている。タグの値4 1 5は同名タグがとるタグの値を数器出力する。

【0 0 1 9】例えば、入力XMLファイル7 1 0中に含まれる、1レコード中に一つだけ存在する「タイプ」タグの数は1 0個なので同名タグ数4 1 3は1 0個で、入力XMLファイル7 1 0中に含まれる「M 0 0 1」を取得し、「基本情報」/「マーク1 1 D」タグ7 2 2を取り出され、「基本情報」/「マーク1 1 D」タグ7 2 3を分割して表示する。タグ7 2 3は分割対象タグではない(ステップ5 0 8)。

【0 0 1 10】図7に入力元XML 1 9 0を表示装置7 0に表示する。なお、タグリスト1 2 0なども外部記憶装置7 0に保持しておけば、あとで再利用が可能である。

【0 0 1 10】以後、ユーザから検索要求や端末要求等があると、処理装置1 0では、主キーインデックス1 5 0をもとに分割対象タグリーフィルタ XML 1 9 0を表示装置7 0に表示し、ユーザが分割対象タグを選択すると、該当分割XMLデータを外部記憶装置7 0から読み出して表示装置2 0に表示する。

【0 0 1 11】図7に入力元XML 1 1 0の具体例(入力XMLデータ1 1 0)を表示し、入力元XMLファイル7 1 0を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 2の値「M 0 0 1」を取得し、「基本情報」/「マーク1 1 D」タグ7 2 2のタグの数を一つ増加する。「基本情報」/「マーク1 1 D」タグ7 2 3は「M 0 0 1」を取得し、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 3は分割対象タグではない(ステップ5 0 8)。

【0 0 1 12】図7に入力元XML 1 1 0の具体例(入力XMLデータ1 1 0)を表示する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 3、「商品名」タグ7 2 4、「基本情報」/「消費電力」タグ7 2 5、「基本情報」/「価格」タグ7 2 6を持つ。タグ7 2 6は子ノードに「値」タグ7 2 7、「出力」タグ7 2 8、「タイプ」タグ7 2 9の順に処理を操作する。タグ7 2 8、「出力」タグ7 2 9の順に処理を操作する。タグ7 2 9の順に処理を操作する。タグ7 2 9の順に処理を操作する。

【0 0 1 13】図3は主キーインデックスXML 1 5 0の具体データXML 1 8 1 0を示す。図9は同一の商品データXML 1 1 0の具体例(分割XMLデータ生成手段1 0)を示す。図10は分割対象タグリーフィルタ XML 1 0 1 0の具体例(分割XMLデータ生成手段1 0)を示す。

【0 0 1 14】タグリスト生成手段1 1 は、まず、一つの入力XMLファイル7 1 0に着目する(ステップ3 0 1)。そして、1レコードを表すタグがまだ存在するか判定する(ステップ3 0 2)。図7に示す入力XMLファイル7 1 0には商品のデータが複数表示されており、ルートタグ「document」の直下に「商品」タグがある。商品タグで囲まれた部分が一つの商品データを表す。そこで、入力XML 7 1 0において1レコード目の商品データ7 2 0を表示する(ステップ5 0 3)。そして、1レコード(1商品データ)中にまだ着目してないタグが存在するか判定する(ステップ5 0 4)。1商品データ7 2 0を表すタグ名、「基本情報」/「マーク1 1 D」タグ7 2 2とタグの値「M 0 0 1」である。「基本情報」/「マーク1 1 D」タグ7 2 3と

【0 0 1 8】図4はタグリストの具体例であり、図7の入力XMLファイル7 1 0の全商品データの解析結果を上記のように仮定して、同名タグの値の多い順、同名タグと同じ値の多い順とタグの値を並べたものである。図4中、4 1 0がタグリストを示す。該データリスト4 1 0は、入力XMLデータ7 2 1中に含まれる11レコード(1商品データ)中に一つだけ出現するタグ名4 1 2が入力XMLデータ7 1 0中に含まれる同名タグ数4 1 3、同タグ名で同じタグの値の数4 1 4の大さい順に順位4 1 1付けして並べリストである。ここでは、同タグ名で同じタグの値をとる数を大きい順に数器目まで示し、タグリスト4 1 0では4番目まで出力されている。タグの値4 1 5は同名タグがとるタグの値を数器出力する。

【0 0 1 9】例えば、入力XMLファイル7 1 0中に含まれる、1レコード中に一つだけ存在する「タイプ」タグの数は1 0個なので同名タグ数4 1 3は1 0個で、入力XMLファイル7 1 0中に含まれる「M 0 0 1」を取得し、「基本情報」/「マーク1 1 D」タグ7 2 3を分割して表示する(ステップ5 0 6)。「基本情報」/「マーク1 1 D」タグ7 2 3の名前とタグの値「M 0 0 1」を保持する(ステップ5 0 7)。「基本情報」/「マーク1 1 D」タグ7 2 3は分割対象タグではない(ステップ5 0 8)。

【0 0 1 10】図7に入力元XML 1 9 0に戻る。商品7 1 0を表示する(ステップ5 0 4)。再びステップ5 0 4に戻る。商品7 1 0を表示するタグ群の中に着目してないタグは存在する。「基本情報」/「マーク1 1 D」タグ7 2 2は主キーを表すタグである(ステップ5 0 6)。「基本情報」/「マーク1 1 D」タグ7 2 2の名前とタグの値「M 0 0 1」を保持する(ステップ5 0 7)。「基本情報」/「マーク1 1 D」タグ7 2 3は分割対象タグではない(ステップ5 0 8)。

【0 0 1 11】図7に入力元XML 1 1 0に戻る。タグ7 2 3は「M 0 0 1」を取得し、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は主キーを表すタグである(ステップ5 0 6)。「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 7)。「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグである(ステップ5 0 8)。そこで、「基本情報」/「マーク1 1 D」タグ7 2 3は分割対象タグではない(ステップ5 0 9)。

【0 0 1 12】図7に入力元XML 1 1 0に戻る。商品データ分割手段1 2は、1商品データ7 2 0を表示する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 3、「商品名」タグ7 2 4、「基本情報」/「消費電力」タグ7 2 5、「基本情報」/「価格」タグ7 2 6を持つ。タグ7 2 6は子ノードに「値」タグ7 2 7、「出力」タグ7 2 8、「タイプ」タグ7 2 9の順に処理を操作する。タグ7 2 8、「出力」タグ7 2 9の順に処理を操作する。タグ7 2 9の順に処理を操作する。

【0 0 1 13】図7に入力元XML 1 1 0に戻る。タグ7 2 3は分割対象タグではない(ステップ5 0 9)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は主キーを表すタグである。タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 10)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 11)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 12)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 13)。

【0 0 1 14】タグリスト生成手段1 2は、まず、入力XMLデータ分剖手段1 2は、1商品データ7 2 0を指定した場合について説明する。

【0 0 1 15】図7にXMLデータ分剖手段1 2の処理フロー1を示す。この結果、入力元XMLファイル7 1 0に含まれる全商品(全レコード)について、1レコード中のタグ群中のタグの出現頻度と同名タグで同じ値の数全てで解析される。

【0 0 1 16】ここで、図7の入力元XMLファイル7 1 0に含まれる全商品データの解析結果は次のようにになったとす

して同じく図4の「基本情報」/「マーク1 1 D」タグを指定した場合について説明する。

【0 0 1 17】ここで、図7の入力元XMLファイル7 1 0に含まれる全商品データの解析結果は次のようにになったとす。この結果、入力元XMLファイル7 1 0に含まれる全商品(全レコード)について、1レコード中のタグ群中のタグの出現頻度と同名タグで同じ値の数全てで解析される。

【0 0 1 18】図3はタグリスト生成手段1 3 0と主キー対象タグ1 4 0を選択する。この分割対象タグリーフィルタ XML 1 3 0と主キー対象タグ1 4 0は記憶装置6 0に格納される。処理装置1 0のXMLデータ分剖手段1 2は、入力元XMLデータ1 1 0を、分割対象タグ1 3 0によって分割して分割XMLデータ1 6 0、1 7 0、1 8 0を生成し、同時に主キー対象タグ1 4 0の値と該当レコードが取り出される。そこで、一つの商品データ(1レコード)7 2 0に着目し(ステップ3 0 3)、その中にまだ取り出していないタグがまだ存在するか判定する(ステップ3 0 4)。1商品データ7 0の値にはまだ取り出していないタグが存在する。そこで、一つの商品データ(1レコード)7 2 1を取り出される。ステップ3 0 5)。そして、この取り出したタグは値を持ち、同階層タグに同名タグが存在するか判定する(ステップ3 0 6)。「基本情報」タグ7 2 1は値を持たない。そこで、「基本情報」タグ7 2 1を取り出される。ステップ3 0 7)。その間にまだ取り出していないタグが存在するか判定する(ステップ3 0 8)。

【0 0 1 19】図4はタグリストの具体例であり、図7では、一つの商品データ(1レコード)7 2 0に着目する。そこで、複数するタグが判定する(ステップ5 0 5)。そして、複数するタグが判定する(ステップ5 0 6)。「基本情報」タグ7 2 1は主キー対象タグではない。そこで、次に着目するタグは同じ値の多い順とタグの値を並べたものである。図4中、4 1 0がタグリストを示す。該データリスト4 1 0は、入力XMLデータ7 2 1中に含まれる11レコード(1商品データ)中に一つだけ出現するタグ名4 1 2が入力XMLデータ7 1 0中に含まれる同名タグ数4 1 3、同タグ名で同じタグの値の数4 1 4の大さい順に順位4 1 1付けして並べリストである。ここでは、同タグ名で同じタグの値をとる数を大きい順に数器目まで示し、タグリスト4 1 0では4番目まで出力されている。タグの値4 1 5は同名タグがとるタグの値を数器出力する。

【0 0 1 20】例えば、入力XMLファイル7 2 3を分割して表示する。タグ7 2 3は分割対象タグではない(ステップ5 0 8)。

【0 0 1 21】図5はXMLデータ分剖手段1 2の処理フロー1を示す。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4が4 0個であるので、同タグが6 0個。「ハンドタイプ」タグ7 2 4が4 0個であるので、同タグ名で同じタグの値4 1 4は1 0個多いのが6 0個、2番目に多いのが4 0個となり、タグの値4 1 5には「床面積タイプ」と「ハンドタイプ」と出力される。

【0 0 1 22】図5はXMLデータ分剖手段1 2の処理フロー2を示す。また、「タイプ」タグ7 2 9が4 0個である。タグ7 2 9は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグ名4 2 0として「基本情報」/「商品名」タグ7 2 4に着目する。「基本情報」/「商品名」タグ7 2 4は主キーを表すタグではない(ステップ5 0 6)。「基本情報」/「商品名」タグ7 2 4は分割対象タグ名4 3 0として「基本情報」/「商品名」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」と「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグである(ステップ5 0 8)。そこで、「基本情報」/「マーク1 1 D」タグ7 2 3は分割対象タグではない(ステップ5 0 9)。

【0 0 1 23】同様にして、1商品データ7 2 0を表示する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 3、「商品名」タグ7 2 4、「基本情報」/「消費電力」タグ7 2 5、「基本情報」/「価格」タグ7 2 6を持つ。タグ7 2 6は子ノードに「値」タグ7 2 7、「出力」タグ7 2 8、「タイプ」タグ7 2 9の順に処理を操作する。タグ7 2 8、「出力」タグ7 2 9の順に処理を操作する。タグ7 2 9の順に処理を操作する。

【0 0 1 24】図5はXMLデータ分剖手段1 2の処理フロー2を示す。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 10)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 11)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 12)。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は「M 0 0 1」を取得する。この結果、「基本情報」/「マーク1 1 D」タグ7 2 4は分割対象タグではない(ステップ5 0 13)。

【0 0 1 25】ここで、図7の入力元XMLファイル7 1 0において1レコード目の商品データ7 2 0が存在する。そこで、まだ着目していないレコードは存在するか判定する(ステップ5 0 14)。

【0 0 1 26】同様に分割対象タグ群についても、ステップ5 0 4～5 0 9に基づいて処理する。残りのタグの中では主キー対象タグと「床面積タイプ」で現れるものは、分割対象タグ7 2 4を保持する(ステップ5 0 9)。

【0 0 1 27】XMLデータ分剖手段1 2は、「基本情報」/「床面積タイプ」と「床面積」の名前と値「床面積」との間に着目してないタグが存在しなかった場合、主キー対象タグ名及びタグの値を主キーインデックス7 2 9の値は「床面積」と「床面積」として表示する。そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在しない場合は、主キー対象タグ7 2 9の値は「床面積」として表示する。

【0 0 1 28】ここで、主キー対象タグ7 2 0が存在する。そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 10)。

【0 0 1 29】ここで、主キーのタグ名、タグの値は「床面積」と「床面積」として取り扱う。入力元XML 7 1 0に着目してない商品データ7 2 0が存在する。そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 11)。

【0 0 1 30】そして、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 12)。

【0 0 1 31】そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 13)。

【0 0 1 32】図3はタグリスト生成手段1 3 0を示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。そして、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 2)。

【0 0 1 33】図3はタグリスト生成手段1 3 0を示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 34】図3はタグリスト生成手段1 3 0を示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 35】図3はタグリスト生成手段1 3 0を示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 36】図3はタグリスト生成手段1 3 0を示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 37】図3はタグリスト生成手段1 3 0を示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 38】ここで、主キーのタグ名、タグの値は「床面積」と「床面積」として取り扱う。入力元XML 7 1 0に着目してない商品データ7 2 0が存在する。そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 11)。

【0 0 1 39】そして、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 12)。

【0 0 1 40】そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 13)。

【0 0 1 41】タグリスト生成手段1 3 0は、まず、一つの入力XMLファイル7 1 0に着目する(ステップ3 0 1)。

【0 0 1 42】そして、1レコードを表すタグがまだ存在するか判定する(ステップ3 0 2)。

【0 0 1 43】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。

【0 0 1 44】タグリスト生成手段1 3 0は、まず、一つの入力XMLファイル7 1 0からタグリストを示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 45】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 5)。そして、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ3 0 6)。

【0 0 1 46】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 7)。その間にまだ取り出していないタグが存在するか判定する(ステップ3 0 8)。

【0 0 1 47】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 9)。

【0 0 1 48】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 10)。

【0 0 1 49】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 11)。

【0 0 1 50】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 12)。

【0 0 1 51】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 13)。

【0 0 1 52】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 14)。

【0 0 1 53】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 15)。

【0 0 1 54】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 16)。

【0 0 1 55】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 17)。

【0 0 1 56】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 18)。

【0 0 1 57】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 19)。

【0 0 1 58】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 20)。

【0 0 1 59】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 21)。

【0 0 1 60】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 22)。

【0 0 1 61】ここで、主キーのタグ名、タグの値は「床面積」と「床面積」として取り扱う。入力元XML 7 1 0に着目してない商品データ7 2 0が存在する。そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 11)。

【0 0 1 62】そして、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 12)。

【0 0 1 63】そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 13)。

【0 0 1 64】タグリスト生成手段1 3 0は、まず、一つの入力XMLファイル7 1 0からタグリストを示す。該データリスト4 1 0は、入力XMLデータ7 1 0において1レコード目の商品データ7 2 0が存在する(ステップ5 0 1)。

【0 0 1 65】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 5)。

【0 0 1 66】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 6)。

【0 0 1 67】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 7)。

【0 0 1 68】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 8)。

【0 0 1 69】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 9)。

【0 0 1 70】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 10)。

【0 0 1 71】ここで、図7の入力元XMLファイル7 1 0において1レコード目の商品データ7 2 0が存在する。そこで、「床面積」と「床面積」との間に着目してないタグが存在するか判定する(ステップ5 0 11)。

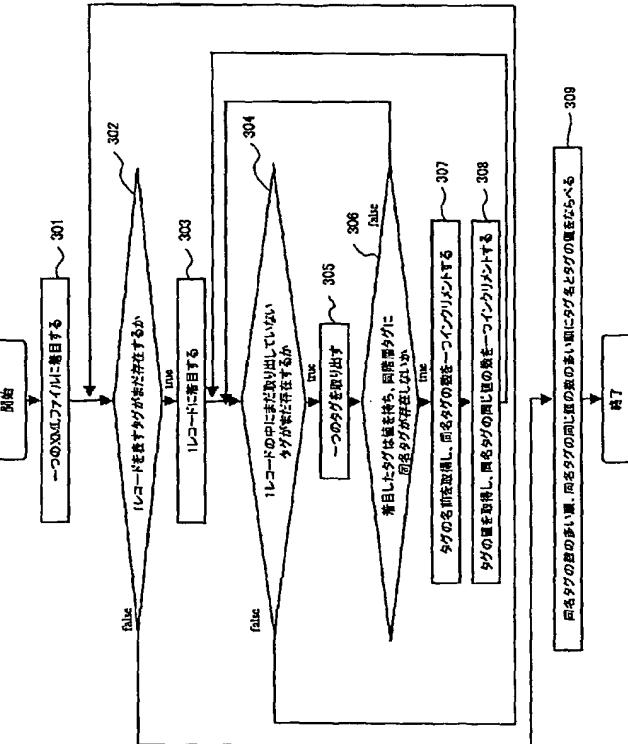
【0 0 1 72】図7に示す入力XMLファイル7 1 0を表示する。タグ7 2 0が取り出される。ステップ3 0 12)。

【0 0 1 73】ここで、主キーのタグ名、タグの値は「床面積」と「床面積」として取り扱う。入力元XML 7 1 0に着目してない商品データ7 2 0が存在する。そこで、「床面積」/「床面積」と「床面積」との間に着目してないタグが

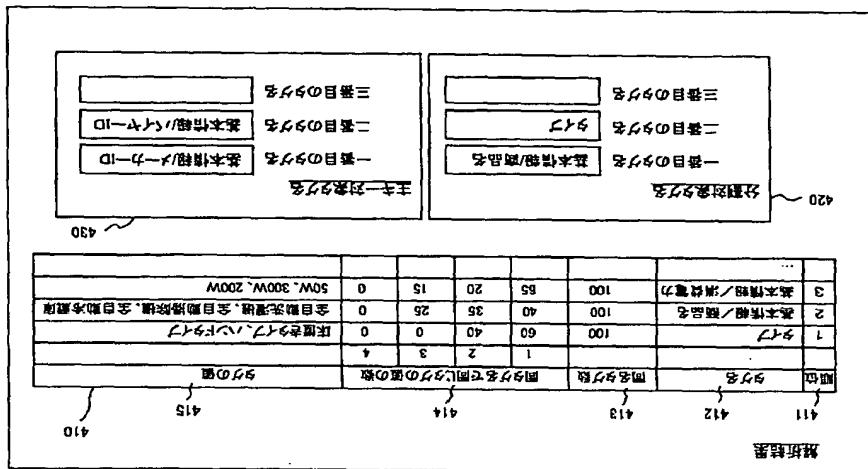


31

## タグリスト生成のフローチャート

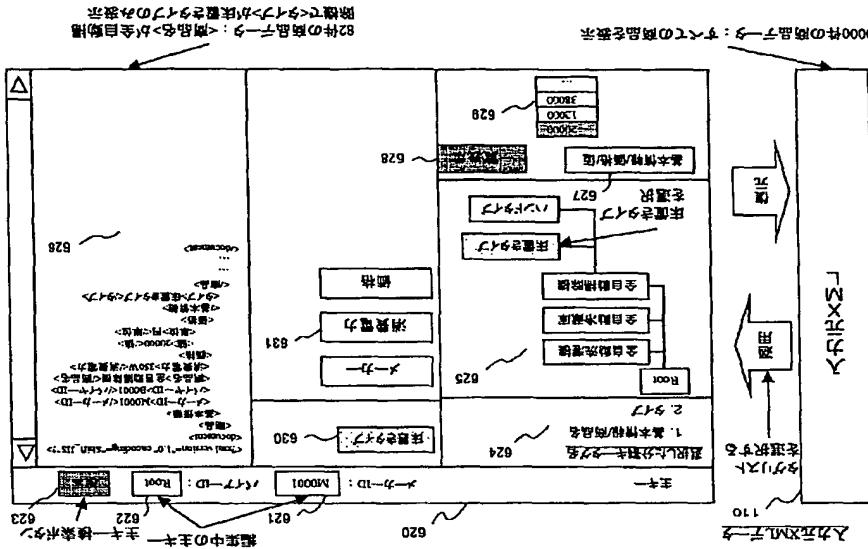


[圖 4]



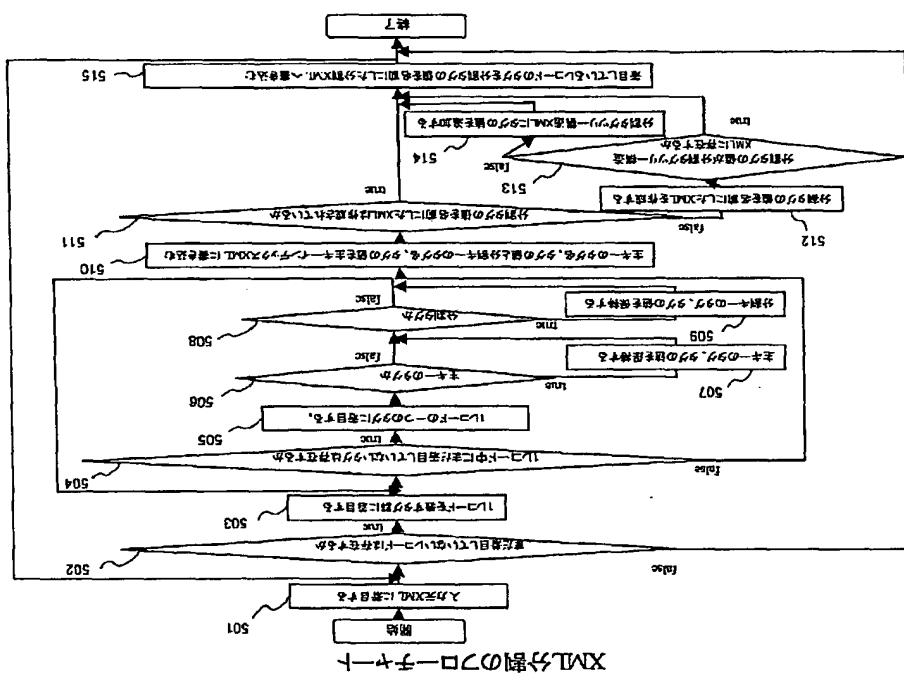
一九一九

6



XML 与一个分割器编辑器连接使用 [Alt+Shift+F11] 快捷键

51



[図7]

## 入力元XMLデータの具体例

```

<?xml version="1.0" encoding="Shift_JIS"?>
<document>
  <商品>
    <基本信息> 721
      <<主力-ID>>M0001<<メーカー-ID>>
      <<バイヤ-ID>>B0001<<バイヤーID>>
      <<商品名>>全自動掃除機<<商品名>>
      <<消費電力>>50W<<消費電力>>
    726 ～ <<価格>>
      <<単位>>円<<単位>> 727
      <<價格>> 20000<<價格>> 728
    <<基本信息>>
      <<タイプ>>床置きタイプ<<タイプ>> 729
    <<商品>>
  <<商品>>
  <<基本信息>>
    <<主力-ID>>M0001<<メーカー-ID>>
    <<バイヤ-ID>>B0002<<バイヤーID>>
    <<商品名>>全自動掃除機<<商品名>>
    <<消費電力>>50W<<消費電力>>
    <<価格>>
      <<単位>>円<<単位>>
      <<價格>> 2000<<價格>>
    <<基本信息>>
      <<タイプ>>床置きタイプ<<タイプ>>
    <<商品>>
  ...
</document>

```

[図8]

## 主キーインデックスXMLの具体例

```

<?xml version="1.0" encoding="Shift_JIS"?>
<document>
  <商品>
    <商品> 820
    <主キー対象> 821
    <基本情報> 823
    <メーカーID>>M0001<<メーカー-ID>>
    <<バイヤ-ID>>B0001<<バイヤーID>> 824
    <<商品名>>全自動掃除機<<商品名>> 825
    <<基本情報>>
      <<主力-ID>>M0001<<主力-ID>>
      <<バイヤ-ID>>B0001<<バイヤーID>> 826
    <主キー対象> 827
    <基本情報> 828
    <<商品名>>全自動掃除機<<商品名>> 829
    <<タイプ>>床置きタイプ<<タイプ>> 829
    <<分割対象>>
      <<商品>>
        <<商品>>
          <<商品>>
            <<商品>>
              <<商品>>
            ...
          ...
        ...
      ...
    ...
  ...
</document>

```

[図8]

## 分割対象タグツリー構造 XML の具体例

```

<?xml version="1.0" encoding="Shift_JIS"?>
<document>
  <全自動洗濯機>
    <床置きタイプ>
      <全自动洗濯機>
      <全自动冷蔵庫>
      <床置きタイプ>
        <全自动冷蔵庫>
        <床置きタイプ>
        <ハンドタイプ>
      ...
    ...
  ...
</document>

```

[図10]

[図9]

## 分割後XML 1の具体例

910

```
<?xml version="1.0" encoding="Shift_JIS"?>
<document>
  <商品>
    <基本情報>
      <メーカーID>M0001</メーカーID>
      <バイヤーID>10001</バイヤーID>
      <商品名>全自動掃除機</商品名>
      <消費電力>50W</消費電力>
      <価格>
        <値>20000</値>
        <単位>円</単位>
      </価格>
    <基本情報>
      <タイプ>床面きタイプ</タイプ>
    <商品>
    ...
  </document>
```